

Exploring Glass as a novel method for hands-free data entry in flexible cystoscopy.

Anonymized for
Submission

ABSTRACT

We present a novel data entry mechanism on Google Glass to annotate cystoscopy findings in a reproducible and hands free manner for use by surgeons during operations in the sterile environment inspired by the current practice of hand-drawn sketches. We developed three data entry variants based on speech and head movements. We assessed the feasibility, benefits and drawbacks of the system with 8 surgeons and Foundation Doctors having up to 30 years' cystoscopy experience at a UK hospital in laboratory trials. We report data entry speed and error rate of input modalities and contrast it with the participants' feedback on their perception of usability, acceptance, and suitability for hospital deployment. The results are supportive of new data entry technologies and point out directions for future improvement of eyewear computers. The findings can be generalised to other endoscopic procedures (e.g. OGD/laryngoscopy) and could be included within hospital IT in the future.

AUTHOR KEYWORDS

Wearable computing; eyewear computing; human-computer interaction; clinical data entry; endoscopy; multi-modal interaction; head gestures; speech recognition; quantitative evaluation; qualitative evaluation.

ACM Classification Keywords

H.1.2 Models and Principles: Human factors; H.5.1 Information interfaces and presentation (e.g., HCI): Artificial, augmented, and virtual realities; H.5.2 Information interfaces and presentation (e.g., HCI): Benchmarking; H.5.2 Information interfaces and presentation (e.g., HCI): Input devices and strategies; H.5.2 Information interfaces and presentation (e.g., HCI): Ergonomics; J.3 Life and Medical Sciences: Medical Information System

INTRODUCTION

With an estimated 151,000 new cases diagnosed in Europe in 2012 and 429,000 new cases diagnosed worldwide in 2012, bladder cancer is one of the most common cancers among both men and women (4th most common among men and 14th most common among women in the UK [1]. In the UK, despite bladder cancer rates dropping by between 25 and 35% since 1990, it remains the 7th most common cancer related cause of death, reflecting both the aggressive nature of the disease and the inherent difficulties

in diagnosis and management. As cigarette smoking is the major risk factor, an increase in incidence in the developing world is expected in the future [2, 3].

The majority of bladder cancer (~80%) is known as Non Muscle Invading Bladder Cancer (NMIBC) and generally presents with painless symptoms, either haematuria (blood in the urine which is not always visible) or urinary symptoms easily attributed to infection such as dysuria (pain while passing urine) [4, 5]. NMIBC ranges from low-grade to high-grade with associated rises in risk of mortality but even slow-growing, low-grade tumours have an extremely high rate of recurrence post treatment of around 31% rising to 78% for high-grade tumours [2, 6]. These recurrences require multiple follow-up procedures and lengthy surveillance over several years.

Late presentation to health care providers, aggressive disease and high recurrence rates combine with technological strategies for diagnosis and management that are expensive and require a high degree of clinical expertise to cause NMIBC to be considered one of the costliest cancers per capita. This is both in terms of financial and human cost; diagnosis, treatment and subsequent long term surveillance can have a substantial and negative impact on quality of life [6-8].

Currently the diagnostic Gold Standard in NMIBC is cystoscopy, the use of a flexible endoscope inserted through the urethra in order to image the urinary tract and bladder as well as perform biopsies to enable grading and staging of cancers found, resection of tumours and introduction of chemotherapeutic agents. This imaging is also used as a reference point for subsequent surgery where necessary. While guidelines issued by clinical bodies such as the UK's National Institute for Clinical Excellence (NICE) [9] and The Royal College of Surgeons (RCS) [10] provide guidance on performing such procedures and what should be recorded (e.g. size, number and appearance of lesions) there is no real standardized, organisational methodology for how this should be logged to best effect. Considering the impact of bladder cancer globally there is a clear need for novel technological solutions to improve any aspect of care but most previous research has been in the field of non-invasive diagnostics by urinary bio-markers, high definition (HD) image capture, optics and improvements in endoscopes rather than in methods of data recording and organisation. Novel methods for optimising data collection

- LICENSE: THE AUTHOR(S) RETAIN COPYRIGHT, BUT ACM RECEIVES AN EXCLUSIVE PUBLICATION

and recording of findings in medical notes therefore offer a useful addition to this field of research.

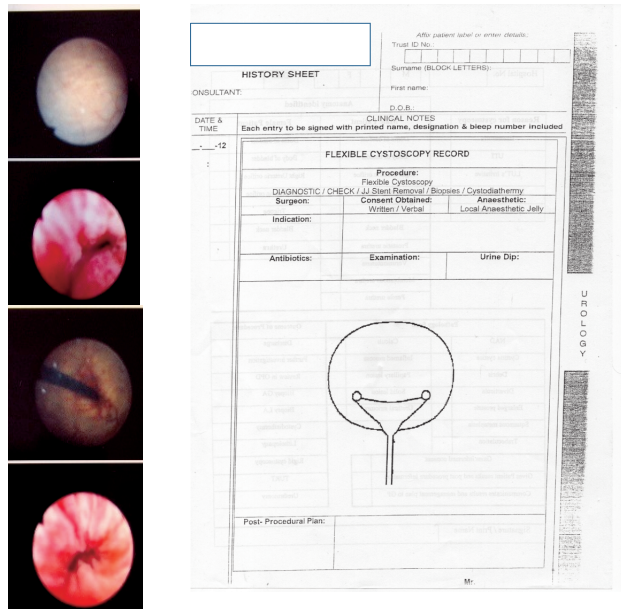


Figure 1. Left: The view through a cystoscope captures only a partial view of the ovoid bladder. Without reference points, it is often not possible to accurately gauge in which direction the cystoscope is oriented: This prevents still images from being used as the sole documentation approach. Right: Example of NHS cystoscopy proforma. A hand-drawn schematic bladder diagram is used to indicate location and appearance of lesions.

We began this project by consulting closely with Urological surgeons of various grades and levels of experience to establish where to concentrate our efforts and what improvements would be most useful to them when making patient notes. We found out that static pictures taken from a cystoscope (*see fig. 1. left*) are felt to be insufficient in documentation of location of tumours due to the lack of reference points with which to orient the viewer. Therefore, hand-drawn bladder sketches are often used, either alongside or instead of the captured image. (*See fig. 1. right*). The issues identified with this approach were inconsistent reporting methods between users, inaccuracy of drawing from memory and high quantity of paperwork per patient leading to overly complicated patient notes.

The key contributions of this paper are:

- Identification of limitation of current data entry mechanism through interviews with surgeons.
- Development of a hands-free data entry system relying on a comparatively low-cost eyewear computer with three input modalities: speech and discrete or continuous head movements.
- Technical and usability evaluation with 8 surgeons and Foundation Doctors.

- Comparative evaluation of learning curve, data entry speed and error rate of each input modality
- Qualitative evaluation through exhaustive interviews and questionnaires of the surgeon's perceptions of usability, acceptance and suitability.

STATE OF THE ART

Flexible cystoscopy is not routinely performed in a sterile environment but involves the use of sterilized equipment that requires highly trained staff using both hands to operate while at the same time interacting with the patient. As both hands are utilized, notes on findings including number, position and appearance of lesion(s) must be recorded after the patient has been reassured, allowed to redress and left the room, therefore relying on the physician's memory.

Recent years have seen rapid advances in the provision of high quality HD still images and video capture to medical practitioners for diagnostic and interventional procedures [11]. These technological advances often come as a 'stack' comprising surgical instruments and software back-end which require the updating of existing hospital equipment leading to high financial investment (cost can be upwards of £50,000 per 'stack') which is not always justifiable when serviceable equipment remains in use. However, it appears that annotated notes and still images remain the preferred method of recording such data during flexible cystoscopy, rather than video capture which involves longer examination times and much larger data storage and processing [12].

A major disadvantage to the current method for gathering data is that of the high number of separate notes and images, recorded by multiple clinicians, within the patient file; not only does this allow for a high degree of user variance in the reporting of what are objective findings (both in terms of appearance and exact location of lesions), the lack of a permanent and persistent annotated image allows for a degree of inconsistency when attempting to track progression of disease as the same clinician may well not be performing subsequent cystoscopy procedures and individualized style and approaches to notation are common. As well as providing an imperfect record of the evolution of lesions over time this can have an impact on the efficiency of analysing public health data. Another disadvantage is that of the reliance on memory, inserting an unnecessary burden of potential error into the procedure.

Computerized endoscopy reporting tools exist [13] and go some way towards integrating paper notes and digital images made but offer little more than is already possible as data is still entered after the fact. The current situation is therefore often a combination of digital still images together with handwritten or typed notes and annotated diagrams, either printed out or stored as part of the EHR or a combination of the two. Potential novel data entry systems could include foot pedals [14], head mounted display (HMD) or wearable technology utilizing speech to text or

gesture based input [15-17], gesture controlled devices [18], incorporating new controls into existing endoscopic equipment [19, 20], or using a 'middle-man' to transcribe notes dictated by the clinician which would require additional highly trained personnel. In applying novel technological solutions to these issues there are multiple points to consider. Foot pedals, although already widely used by surgeons in theatres, have several drawbacks. They are generally used for binary actions such as activating tools and even then suffer from issues of accuracy and poor ergonomics [14, 19, 20]. They would also be inappropriate for text based input requiring increased time and coordination for accuracy and an additional screen on which to view the text input, adding a second or third screen to an already ergonomically poor environment [21]. Wearable gesture controlled devices have shown potential in data entry and in controlling external devices, as have ambient gesture sensors, but text input is limited and too slow to be used in these circumstances [22-25]. Adding 'keyboard' controls to existing endoscopic equipment risks over-complicating the interaction with the device. While input speed could approach normal typing speed with a device of this sort [26], control of the endoscope could be impaired.

HMDs and eyewear computers though, appear to overcome some of these issues. Being head-mounted they should not interfere with fine motor control of the endoscope nor have an appreciable impact on infection control. They reduce the ergonomic load of further screens and could combine gesture control and speech to text input for data entry and device control. Google Glass is a much mediated example of a growing trend towards commercial eyewear computers (e.g. Vuzix, Sony SmartEyeGlass). Glass is designed to be worn like conventional glasses. It includes a computerized central processing unit, an integrated micro-display screen, high-definition camera, bone-conduction sound transducer, wireless connectivity and different types of sensors: a microphone for ambient sound recording and speech recognition, an inertial measurement unit to capture head movements and a capacitive touchpad. Interaction with Glass is mostly through touch gestures on the capacitive pad, but the sensing capabilities of the device allow a potential hands-free mode of operations. The micro-display screen presents information with minimum intrusiveness. Many of these characteristics are representative of what eyewear computers currently on the market offer. With these characteristics, Glass garnered interest as a tool in the workplace and in the medical sector.

Medical research involving Google Glass and similar devices has concentrated on image capture/video recording capabilities and wireless communication, using these to enhance teaching [27-31], to provide offsite consultations [32-34], to refer to diagnostic imaging or patient data during surgical or interventional procedures [35-37] and to record medical data [32, 38, 39].

Outside of medicine, HMDs and eyewear computers have been investigated for use in on-site inspection utilizing both data retrieval and recording [40], for image capture and environmental analysis [41], data recording [42], for capturing and guiding lab experiments [43, 44], enhancing teaching [45], guidance systems for manual assembly and repair tasks [15, 46], control of drones [47] and as a replacement for mobile technology for those with physical disability that complicates or impedes use with hands [48]. Google Glass, and similar wearable devices, clearly open up a new dimension for hands free data entry and documentation. Furthermore, current eyewear computers may be up to two orders of magnitude cheaper than a full-fledged computerized endoscopy 'stack'. Therefore, HMDs and eyewear computers could provide a low-cost route to standardizing notation to augment digital visual data captured by endoscopes and improve information feedback to surgeons planning surgical procedures.

EXPERIMENTAL SETUP

Initial requirements and design input

As eyewear computers can be used hands free, which is a major consideration in clinical environments where infection control is of high importance, we wished to investigate how a device such as Glass could be used to augment current data entry approaches relying on hand-drawn sketches. We held multiple discussions with 3 consultant Urologists and 4 surgical Registrars at a UK based Hospital Trust along with observations of how cystoscopy is performed, in order to determine areas which staff felt could be improved by the addition of a novel HMD. In these sessions the basic capabilities of Glass were explained and ideas regarding improvements to hand drawn images were related to how computerized endoscopy reporting was currently experienced. A key finding was concern over lack of consistency in reporting results due to inter-user variance. This was identified as being a factor of differences in annotation style and results from multiple cystoscopies being recorded separately which can not only make reviewing patient notes complicated but also potentially generates an enormous amount of extraneous paperwork. Through these discussions it emerged that a diagrammatic representation of the bladder was found to be useful as an adjunct to still images recorded and as a quick reference for consultations but that improvements were possible in the implementation of such a system. As there were also concerns raised regarding the accuracy of the drawings an idea was put forward that a grid overlaying a basic image of a bladder could be designed together with standardized symbols for notation. To address the issue of consistency it was decided that a persistent image, annotated in a standardized manner, could be assigned to each patient when first encountered. This image could then be added to during subsequent follow up providing a single reference point with clear notation to aid in planning surgery, possibly reducing the length of time of surgical

investigations, and aiding in future medical and surgical management.

Based on this initial consideration we set out to assess how to best enter annotations on the schematic bladder using the hands free capabilities supported by Glass.

Bladder Annotation on Google Glass

Based on the design input we developed a digital replica of the current procedure of annotating a hand-drawn bladder sketch. Therefore, the main UI element is a schematic bladder drawing, where the surgeon can pinpoint a location to annotate with an icon representing the nature of the lesion. The application has been implemented on Android Studio using the Glass Development Kit (GDK) following an "immersion" design pattern. This is a Glass-specific design pattern where the application defines its own user interface and takes full control of the user experience.

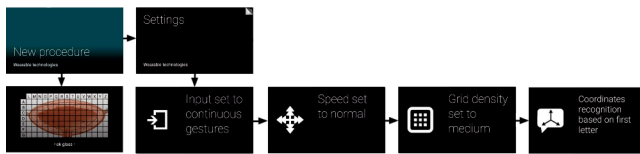


Figure 2. Application navigation flow. The main timeline comprises two cards: new procedure and settings. The settings are composed by four cards, one per parameter.

The application settings offer three hands-free modalities with which to choose a grid cell to annotate: "speech", "discrete head movements" and "continuous head movements".

After launching the application, the user is directed to a start screen. From the start screen (*figure 2*) it is possible to start a new cystoscopy procedure, which allows surgeons to annotate the bladder image, or to move to the settings card to configure the application.

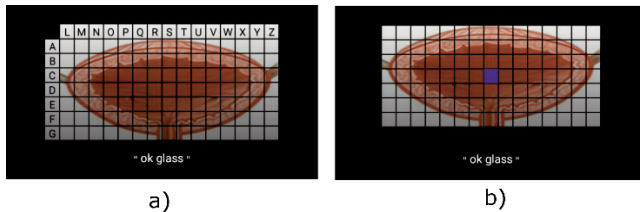


Figure 3. Depending on the input modality the procedure grid shows: a) rows and columns identified by indexes for the speech input modality, or b) a cursor highlighting the cell selected for the input mechanisms based on head movements.

When a procedure starts a new card is loaded, showing a basic image of a bladder. In order to help with data entry, a grid is overlaid on top. It serves as a reference to enter the data observed during the procedure.

In order to populate the grid, the user has to choose each cell individually and indicate the type of annotation to place at that coordinate. The procedure card includes a voice triggered menu, activated when speaking the command "ok glass", which contains a list with the voice-activated

commands required to add or remove annotations. The user then selects the cell (either through speech or head movement mode) and finally enters a number from 1 to 5 corresponding to an icon (e.g. a type of lesion) to insert in the cell.

In 'speech' mode the cell is selected by means of two coordinates (row and column) spoken by the user using the NATO phonetic alphabet (see Figure 3a). The Glass speech recognizer processes the voice input and returns the transcribed text to be used as indexes for the grid. In speech mode an additional setting defines how the captured speech is transcribed into coordinates. It can be set up 'Full match', where the text spoken by the user must exactly correspond to code words in the NATO phonetic alphabet. However, preliminary trials showed that speech recognition errors were frequent with single word recognition. Therefore, the other options are 'First letter', where the first letter of the words spoken by the user is matched against the phonetic alphabet and taken as cell indexes, and 'Similarity', where the words in the phonetic alphabet, which is closer to the words recognized from the user, are taken as indexes. Hereafter we used the 'first letter' mode.

In 'discrete movements' mode, a cursor indicating the cell selected is overlaid to the grid interface (see Figure 3b). The user can move the cursor within the grid using head nods vertically and laterally. The gyroscope is employed to calculate the angular velocity of the device. Those rotational changes with a velocity above a threshold are detected as motion events.

In 'continuous movements', the selected cell overlaid on the grid is moved by using the head as a pointing device. The gyroscope is employed to calculate the head orientation compared to the initial orientation when the procedure was started. The difference in orientation is translated into a two-dimensional motion of the selected cell on the grid. The sensitivity can be configured from the application. Here we used a 'normal' setting where a head rotation of ~60 degrees horizontally and ~35 degrees vertically allows navigation from one edge to the other on the grid.

The Glass application continuously recorded how the user interacted with it. In particular, each cystoscopic procedure leads to a log file. The log file contains timestamped information about each interaction with the device (cell movement direction, speech recognizer output, etc.). This allowed us to obtain precise timings for subsequent analyses.

Protocol

The experimental protocol (see fig. 5) was designed to collect both quantitative and qualitative data about the suitability of Glass to capture findings in cystoscopy. As for all participants this would be the first time they would employ Glass and this specific application, we designed the protocol to assess elements such as learning curve, error rate and accuracy of data entry through the objective logs recorded by the application. We set-up questionnaires and

elicited discussion to compare and contrast these objective findings with the subjective perception of the clinicians regarding usability, clinical deployment, and potential benefits to working practices and patients.

First, subjects completed an informed consent form. Then they were asked to complete a pre-experimental questionnaire that consisted of basic demographics, familiarity with cystoscopy and digital mobile devices and opinions regarding the existing (paper-based and image or video capture) methods of data capture. Afterwards, participants watched a short tutorial video developed by us demonstrating the basic operation of Glass, how to navigate the application and how to use the three input modalities.

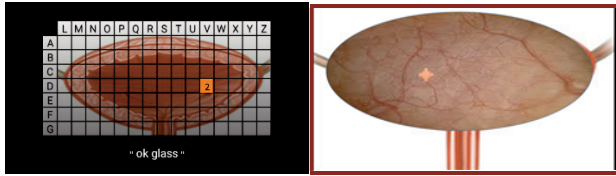


Figure 4. Example data entry instances presented on the laptop with (for training) and without (for deployment) grid overlay.

Participants were then given Glass to wear and asked to adjust it to their comfort before being shown a series of 20 data entry instances for each modality (hereafter called *training phase*). Each data entry instance was shown on a laptop at around 1m from the participants. The screen showed the same view as the Glass application presents (i.e. bladder schema with the grid overlay), with the addition of one cell that is highlighted and showing a number (fig. 4, left). Participants were then asked to navigate to the corresponding grid coordinate and input data that matched the number shown, with all data being recorded by Glass. This included time of starting each task, number of attempts, failed data input and successful data input. Number of cells navigated through was also recorded for ‘discrete’ and ‘continuous’ modalities. Subjects were then asked to complete 5 more data entry instances for each modality without the grid, in order to test the accuracy of data entry in more realistic conditions (hereafter called *deployment phase*), as cystoscopes do not have a grid overlay (fig 4, right).

Finally, participants were asked to complete a post-experimental questionnaire. It aimed at eliciting opinions on how fast volunteers were able to learn how to use both Glass and the application and at comparing all three modalities of data entry. Questions covered, for all three input modalities, perception of speed, accuracy, ease of use, acceptability in the clinical environment to clinicians and patients and for fatigue/discomfort experienced. Questions were then asked to gauge the opinion of the extent such an application could improve current methods of data capture and recording and how it could be integrated into current guidelines. During the completion of this questionnaire verbal feedback was also elicited. The entire experimental protocol was videotaped for further analyses. The

experiment leader was both experienced on the technological side and on the clinical side, and in particular had experience with cystoscopy which allowed framing of the discussion during debriefing. The entire experimental protocol lasted about 90 minutes. Details about the cohort are provided in table 1. Glass was connected to the hospital Wi-Fi network for speech recognition.

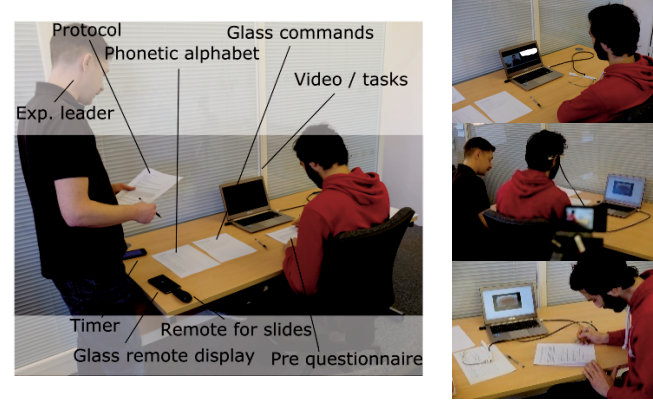


Figure 5. The experimental protocol comprises a pre-experiment questionnaire (left), an introduction video (right, top), the data entry tasks (right, middle) and a post-experiment questionnaire. All experiments are video-taped with the experiment leader taking notes (right, middle). A mobile phone mirrors the Glass screen to check the progress of the task; a remote control is used to switch to the next task when a task is completed. A phonetic alphabet and a summary of Glass commands are provided to the user (left).

User	Experience in urology	Use of vision aids	Experience in mobile technology	Experience of wearables
0	>10	Prescription lenses	Daily use	Fitness tracker
1	<5	Prescription contacts	High daily use	None
2	5-10	None	High daily use	None
3	>10	Prescription lenses	High daily use	None
4	<2	None	High daily use	None
5	<5	Prescription glasses	Daily use	None
6	<2	None	High daily use	None
7	5-10	None	High daily use	None

Table 1. The cohort comprised 8 subjects of various expertise level performing cystoscopy procedures in a UK hospital.

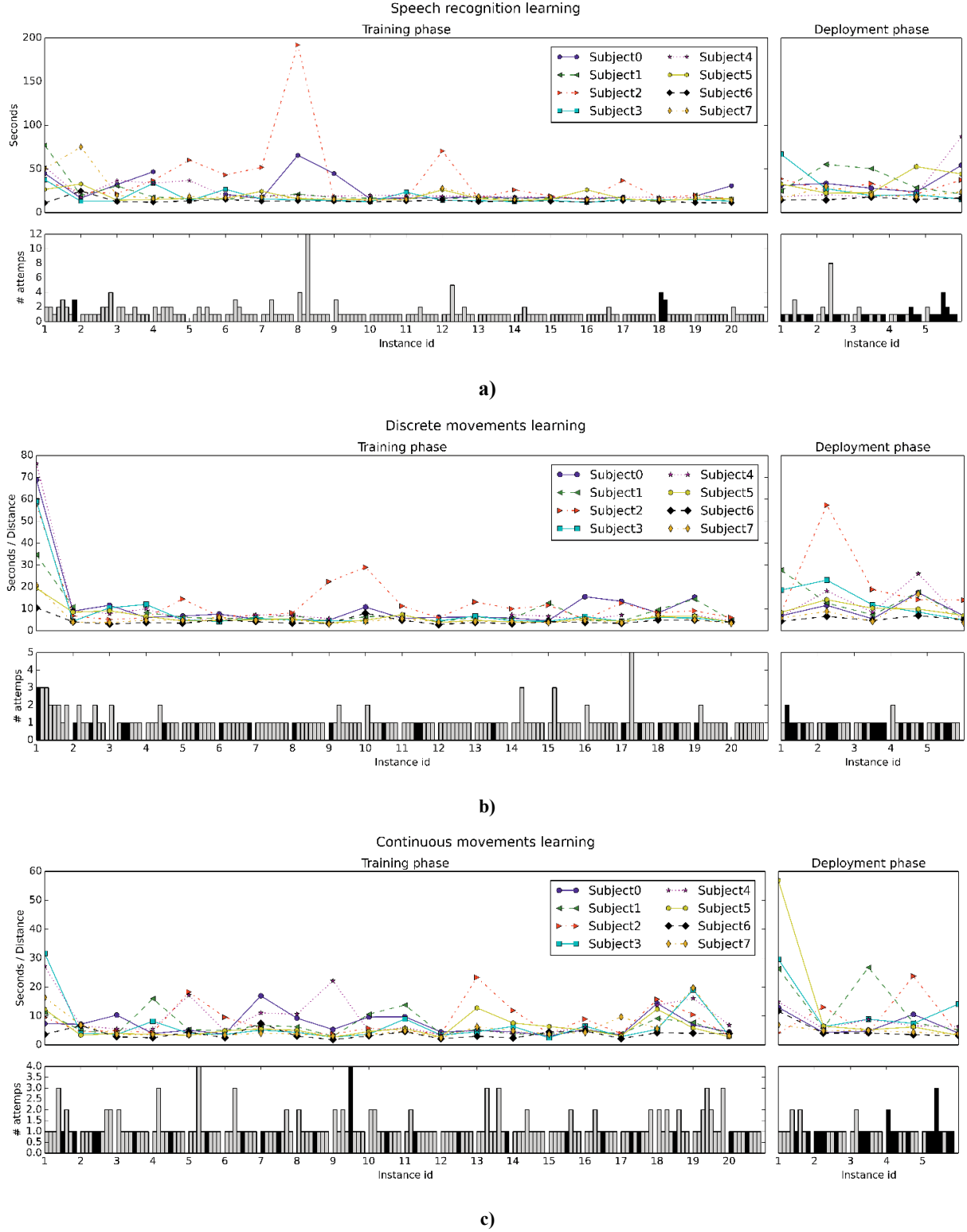


Figure 6: Time (upper plot) and number of attempts per user (lower bars) when using the speech input modality (a), discrete head movements (b) and continuous head movements (c). The groups of 8 bars per instance indicate how many attempts each of the 8 users made at solving the data entry task. If the bar is white, it means that the task was completed successfully, possibly after several attempts. If the bar is black, the user ultimately failed to enter the data correctly. On the left the data is shown for the 20 training instances, and on the right for the 5 instances in the deployment phase.

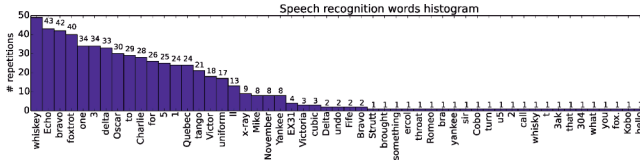


Figure 7. Histogram with the number of identifications per word using the speech recognition modality.

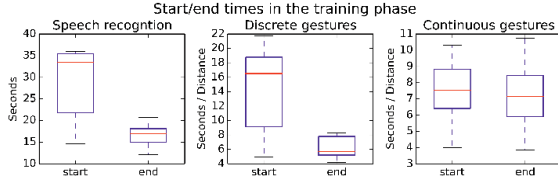


Figure 8. Average data entry time per modality for the first and last 5 instances during the training phase.



Figure 9. Average data entry time per modality for all the instances during training and deployment phases.

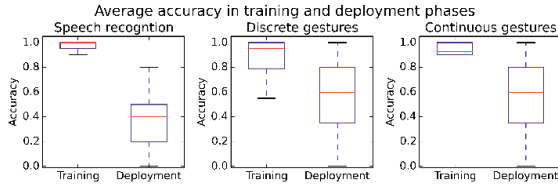


Figure 10. Average accuracy per modality for all the instances during training and deployment phases.

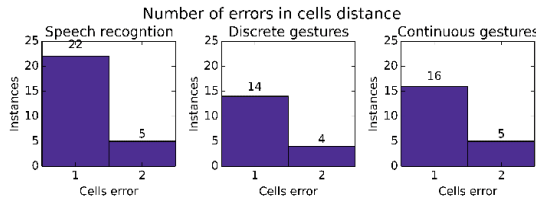


Figure 11. Error in terms of cells distance during deployment phase.

RESULTS AND ANALYSIS

Speed, learning effect and accuracy of entry

Figure 6 reports the time and number of data entry attempts taken to complete the data entry instances in training and deployment phases, for each input modality. The time is given in seconds for the speech modality and in seconds divided by the distance from the centre to target cell in the head movement modalities to normalise for the travel distance. Some users appear as outliers, such as subject 2. In this case, the participant provided continuous verbal

feedback during the actual task, even though all participants were instructed to keep feedback for the end.

The number of data entry attempts was generally higher at the start of the training phased compared to the end, reflecting increased proficiency at data entry, with the exception of the 'continuous' modality, where a "lock on" issue (discussed in "Issues Encountered") explains why the number of attempts does not decrease over time.

Participants became significantly faster over time for the speech and discrete movement modalities, but not for the continuous movement modality (fig 8). Compounding this, continuous movements was about twice as fast as discrete movements at the start of the training phase, while at the end of the training phase the speed of discrete was slightly faster (fig 8). Hence discrete movements required training to increase proficiency, while continuous movements allowed close to optimal speed from the start.

When removing the reference grid in the deployment phase (to better reflect the real condition where the cystoscope does not present such a grid) the time to enter data remained identical for discrete and continuous movements, but increased by almost 50% for speech recognition (fig 9). This is likely explained by the fact that the cell within Glass can be placed using a visual intuition as to what is closest to the cell highlighted on the laptop. However, when using speech recognition participants had to first visually map where the cell shown on the laptop would be within the grid shown in glass, and only then could say its coordinates.

The accuracy of data entry was very high in the training phase, as participants could undo a cell entry and retry until they were satisfied (which is reflected by more time/attempts to complete the task), however accuracy decreased in the deployment phase (fig 10). The accuracy in the deployment phase compares what is the true coordinate of the highlighted cell in the grid-less laptop display with the actual cell participants entered in Glass. Even though the pictures shown on Glass and on the laptop looked different (i.e. more representative of a real cystoscopic procedure), we found out that participants entered a lesion at most 2 cells away from the true location, with the majority of lesions entered exactly 1 cell away from the true location (fig 11). This actually indicates a very low inter-user variability in the pinpointing of lesion location, which is important for a reproducible data entry mechanism.

Speech recognition posed issues as single word recognition was not sufficiently accurate. Figure 7 shows a histogram of recognised word occurrence. Ideally, all the words should belong to the phonetic alphabet and numbers 1 to 5, but the long tail indicates misrecognised words as well as user errors (e.g. 'hello', where the user likely was confused about the data entry procedure). As some errors were frequent (e.g. 'fife' instead of 'five') the application that was deployed included a substitution table to correct common recognition errors (e.g. 'II' instead of '2', 'for' instead of '4').

Current System of Annotation (paper based)

Attitudes towards the accuracy and consistency of the current paper-based system of using multiple drawings to identify, describe and track lesions within patient notes were elicited. The highest scores on a 5 point scale of *Poor* to *Excellent* were merely satisfactory, very much reflecting the issues with current methodology identified in the genesis of this project. 6 out of 8 participants rated the drawings as either *Poor* or *unsatisfactory* on a 5 point scale from *Poor* to *Excellent* when asked about the accuracy and consistency of tracking lesions over time by the current paper based method, with two rating the current system as poor or unsatisfactory on all measures (accuracy and consistency of identifying lesions, tracking lesions over time and locating lesions). 3 out of 5 of the low scores in this regard came from more junior clinicians and two of the more senior clinicians rated the current methods as poor or unsatisfactory across all measures.

Preferred Glass Entry Modality

When asked to rank the different data entry modalities in order of preference ‘speech’ was the most popular (4 most preferred, 3 second choice), closely followed by ‘continuous’ (3 most preferred, 3 second choice) and with ‘discrete’ being by far the least preferred (1 most preferred, 2 second choice). Commenting on this, speed and accuracy of data entry were cited, but ‘speech’ and ‘continuous’ were also described as being more intuitive to learn (3 participants cited this specifically in their comments). This is borne out in the quantitative data recorded by Glass, with ‘continuous’ modality exhibiting the fastest data input times and being very consistent in this throughout the testing. ‘Speech’ input was the slowest in terms of speed of data entry at the beginning of testing and despite an average 42.2% decrease in data entry time by the end of testing was still the slowest data entry modality. ‘Speech’ was, however, the most accurate of the 3 modalities during training. Learning time for the ‘speech’ modality may therefore be a factor; issues identified from the comments of participants were the experience of phonetic alphabets, accent (one subject did not use English as their first language) and inflection. It was noticed that a ‘High Rising Terminal’ inflection, common among American English speakers, showed some success in improving voice recognition. ‘Discrete’ was only the first preference of one user and was generally described as “*Clumsy and slow*”, “*Jerky*”, “*Uncomfortable*”, “*requiring overly sharp head movements*” or “*hard to use*” in the comments provided. Perceptions regarding the accuracy or speed of this modality did not match the data recorded, with ‘discrete’ being the second fastest in terms of data entry at the start of training and exhibiting an average 55% decrease in data input time making it the fastest data entry modality by the end of training. ‘Discrete’ was also more accurate than ‘continuous’ but the physical discomfort experienced by all but one user must be accounted for, especially if a device such as this were to be used for prolonged periods. The

difference in accuracy between ‘discrete’ and ‘continuous’ may be explained by the “Lock on” issue whereby users described problems with the selected cell moving while using voice commands to enter the correct icon (see below). It is possible that addressing this issue could further decrease data entry times for ‘continuous’ gestures.

Potential to Improve Data Recording

When asked about the system’s “*potential to improve the accuracy of describing the site of bladder lesions*” and “*potential to improve the accuracy of tracking progression of bladder lesions*” the response was positive, with 7 out of 8 participants scoring 4 or 5 (on a 5 point scale from *Not at all* to *Considerably*) for both measures. This reflects a high satisfaction with utilising a digital method of data collection and is a positive indicator of the potential for hands free use of HMDs in clinical environments. The one user that felt the system to be poor had a visual field defect.

The perception of the potential of an application of this sort to improve the inter-user variability noted with the current method of data collection (whereby multiple clinicians perform subsequent cystoscopies using individualized annotation styles) was also high with 7 out of 8 participants scoring 4 or 5 (on a 5 point scale from *Not at all* to *Considerably*). While the potential to “*make an improvement in how bladder lesions are described*” was also rated as high (6 out of 8 participants scored 4 or 5 on a 5 point scale from *Not at all* to *Considerably*), “*potential to improve the accuracy of diagnosis of bladder lesions*” scored lower, with 5 out of 8 participants scoring 3 out of 5 on the same 5 point score. This could be explained by the fact that diagnosis is more often made on the basis of biopsies taken from the site of lesions once clinical suspicion of a cancer has been confirmed by the cystoscopy findings. This means that the notes taken during cystoscopy are used more often for management decisions such as planning further intervention; this is borne out by the high scores regarding the system’s “*potential to improve patient management*” and “*communication of diagnosis to patients*” (5 out of 8 subjects rated 4 or 5 on the same 5 point score). When asked about the system’s “*potential to speed up the process of note taking*” 4 participants scored 4 or 5 on the same scale and when asked “*do you feel that a persistent and annotatable image such as that presented could effectively reduce paperwork*” 7 out of 8 subjects scored 4 or 5 on the same scale.

Acceptability in Clinical Environments

The questionnaire assessed acceptability to practitioners and to patients. Acceptability of a HMD used in this manner was high. When asked “*how easy would you find the integration of Glass into your own practice*” 5 out of 8 participants scored 4 or 5 on a 5 point scale from *Very Difficult* to *Very Easy*. Only 2 out of 8 users felt that it could present issues; importantly, physical issues with using the Glass itself were mentioned by both. When asked “*How acceptable would you find the use of a device such as Glass*

during cystoscopy” 6 out of 8 participants scored 4 or 5 on a 5 point scale from *Entirely Unacceptable*” to *“Entirely Acceptable”* Interestingly none of the users predicted that Glass would interfere with hand eye coordination or their use of other equipment while working and all felt that Glass would have little or no impact on infection control.

“Acceptability to patients” was estimated by the subjects to be high with 6 out of 8 subjects scoring 4 or 5 on the same 5 point scale, although there were concerns raised by 3 participants regarding the potential impact on Doctor-patient interaction of introducing a piece of equipment that *“Sits between me and the patient”* and could present distractions by *“Spending more time looking at the Glass than the patient”*. Balancing this were the views of 2 participants who felt that it could have an entirely positive impact on interaction, scoring 5/5 in this regard.

Issues Encountered

The most common problem encountered was the inability to *“Lock on”* to cells while using ‘continuous’ modality (all 8 participants mentioned this). The current Glass GDK does not notify applications when “ok glass” is detected. It only sends an event when a menu command is detected after that. Consequently, the application is unaware of when the user decides to “lock on” to a cell by saying “ok glass”; it only knows when “new cell” is spoken. In the interval between “ok glass” and “new cell” the head tracking algorithm keeps running, which requires users to keep their head still until they pronounce “new cell”. This often meant that the cursor would move from the selected cell while voice commands were being given, resulting in errors and accounting for the majority of repeated attempts. Despite this, ‘continuous’ was very popular as a data entry method and this represents a major factor when considering the design of such software. Comfort and perception of usability of ‘continuous’ was high, being the second most preferred modality (see above) and this should be weighed against actual efficiency. It would be preferable however to address this issue in the future. Another major issue encountered was that of the speed of voice recognition. 5 subjects became frustrated at delays between enunciation of a command and its recognition by Glass and comments such as *“This is not going to be fast enough for Doctors”* were offered by 3 subjects. Indeed, the same 5 users raised concerns that Glass itself was not yet responsive enough (in terms of transitioning between screens and responding to voice commands) to introduce into clinical environments.

Three participants also suggested that more than one 2D image could be utilised to increase the accuracy of siting lesions, therefore helping better track lesions over time and as an aid to digital still images recorded by the cystoscope to provide surgeons with more information when planning procedures. 4 participants felt that an HMD could also be useful for other tasks, especially in recording notes and accessing patient data/images during procedures and that an application such as this could work very well in conjunction

with images recorded by the endoscope by giving more reference points for interpreting the images. It was also mentioned by one of the volunteers that the ability to *“Map or merge the Glass image onto the cystoscope image”* would be beneficial.

Glass

When asked about Glass itself, there were issues identified with wearing prescription lenses; Glass in the form that we tested does not have a facility for combining with prescription lenses and so it was sometimes difficult to prevent Glass from moving from an ideal position when resting on top of prescription frames. Those participants using contact lenses reported no issues. One participant had a visual field defect in the eye over which the Glass optic was positioned and this caused a substantial amount of discomfort and difficulty in data entry. This user was concerned about the inability to adjust Glass to use with his other eye and provided the lowest scores across all questions focusing on the system’s ability to improve current note making methods. This user also reported migraine symptoms the following day which they associated with visual issues and neck discomfort from the testing; these are extremely important factors to consider in planning future introduction of such technology into work environments where long usage times and imperfect ergonomics would further exacerbate any physical effects on users. Transient and minor neck discomfort was associated with both ‘continuous’ and ‘discrete’ input by 5 users, citing unnatural neck positions or keeping the neck extremely still to avoid moving the cursor inadvertently. One other participant also raised physical issues with Glass; being particularly small in stature they were unable to position the Glass securely on their head and found that it moved out of place during testing. Interestingly this volunteer provided the second lowest scores in terms of accuracy and potential to improve current note taking suggesting that physical difficulties in using Glass may have a substantial and negative impact on perception of function and user acceptance. One other issue identified was that of the environment in which Glass is to be used, with 3 participants mentioning that a blank white wall as a backdrop to patient encounters would be the only way they could use Glass effectively. They indicated that any variation in the backdrop caused difficulty in visualising details and between normal vision and that which appeared in the Glass optic.

General Discussion

There was an overall positive reaction towards using an HMD such as Glass within daily clinical activities and its potential impact on current note taking within the NHS (UK’s National Health Service) Trust it was trialled in. While some of the issues that we aimed to address with Glass such as standardization of annotation of cystoscopy findings are also targeted by more modern endoscopic technology, there remains a consensus of a continued need for annotated diagrams to be used as an adjunct to HD

endoscopic images in the case of cystoscopy; the spherical nature of the bladder in particular often causes difficulty in locating lesions due to lack of reference points. The ability to make notes of this sort during the procedure rather than as an afterthought was considered a positive addition. Cost is another area where HMDs such as Glass could be beneficial when compared with updating whole endoscopy stacks (estimated at £50,000-100,000 per device). Glass could offer an affordable interim measure to improve notation of findings while also potentially being compatible with future hardware updates due to the universality of its software.

Glass is also suitable for a developing approach to working known as 'bring your own device' (BYOD) whereby employees conduct official work on privately owned technology. This approach allows for very flexible working patterns and is already common among clinicians as regards other medical equipment and computing devices, with stringent guidelines for data protection already in place.

Limitations to this study include its size, with only 8 subjects it makes it difficult to estimate general acceptability across the NHS although the generally positive comments are encouraging and bear more investigation. While the study size was small it did include Surgeons and Junior Doctors across a range of grades, with 7 out of 8 subjects having experience of other NHS Trusts and 4 subjects having experience of working in Private medicine. The positive reactions from those practitioners having experience with other, more modern, equipment suggest that this system could be suitably integrated to other environments. The time that each participant was able to use Glass was also constrained as we reduced the extent of the testing (to under 90 minutes) originally planned in order to accommodate the busy working days of the subjects. It was also not possible to trial Glass with patients yet, as this work is an early acceptability study.

A drawback to speech recognition technology as a whole was encountered whereby any word within Google's dictionary is recognised and the API does not allow limiting words to a specific dictionary (such as e.g. the phonetic alphabet). Any further development of this application for more extensive notation/annotation would necessitate extensive research into commonly used words or phrases for describing findings. Ideally, it would be possible to make available standardized statements to apply to the image or to refine the speech dictionary.

Other input modalities may become available in future eyewear computers, such as eye tracking. They have not been considered in this project as one particular challenge in cystoscopy is the spherical nature of the bladder. It is not possible to capture the entirety of the bladder in a single image of an endoscope and therefore it is not possible to rely solely on gaze direction to point to locations of lesions, as the knowledge of the orientation of the cystoscope within the bladder is missing. However, future cystoscopes may be

equipped with inertial tracking systems, e.g. to monitor surgeon's skills [49] and this could be used to address this registration problem.

CONCLUSION

This paper examined the potential of Glass to provide an affordable, hands free data entry device for use by Urologists to augment the HD images recorded by modern endoscope technology when investigating suspected bladder cancer. To our knowledge there have been no previous studies undertaken investigating the use of HMDs for this purpose specifically. It contributes by showing one approach to systematic lesion annotation, supported by a reference grid, and it provides valuable information on how head gesture sensing and speech recognition is best used in this environment while also presenting user feedback from clinicians working with these procedures on a daily basis.

The feedback from participants was positive suggesting a high acceptability of using HMDs in the clinical environment, with negative feedback focusing on physical issues such as visual defects and the form factor of Glass itself. While the potential for HMDs to improve current note taking systems in terms of consistency and in accuracy of locating and of tracking lesions over time were both viewed positively, as was the potential to reduce paperwork for clinicians, it was felt that the technology was not yet fast or responsive enough to deploy in its current form. The approach to the UI of providing an image with a grid on which to annotate findings was viewed as a positive contribution. Opinions on acceptability to patients were more mixed, with some concerns raised over the impact on communication with patients but also the potential to improve communication of diagnoses. While it was felt that Glass would have little or no benefit in terms of increasing the speed of diagnosis it could make surveillance of cancer patients more accurate and could also aid in providing more accurate data to surgeons. Glass was considered by all users to be easily adaptable to infection control guidelines.

'Speech' and 'continuous' were the preferred data entry modalities, rated mainly on ease of use and speed. This was not represented in the data recorded from Glass which showed 'continuous' and 'discrete' as being the fastest and 'speech' least prone to error of the modalities. 'Discrete,' however, was disliked by the majority for being uncomfortable and difficult to use. Comfort and intuitive learning appear to have a very strong influence on perception of speed and accuracy of function.

Future studies could test the system's impact on ability to use precision equipment and its acceptability in a 'live' clinical environment. Side by side comparison with handwritten drawings would be required to firmly establish the positive impact of a persistent single annotatable image. Further refinement of the application could include audio notes added to images, compatibility with EHR and medical terminology for speech recognition. Finally, other HMDs could be compared to establish ideal ergonomics.

REFERENCES

1. Cancer Research UK. Bladder cancer statistics. Retrieved March 03, 2016 from <http://www.cancerresearchuk.org/health-professiona/cancer-statistics/statistics-by-cancer-type/bladder-cancer#heading-Zero>.
2. C. S.D. Lee, C. Y. Yoon, J. A. Witjes. 2008. The past, present and future of cystoscopy: the fusion of cystoscopy and novel imaging technology. *BJU International* 102, 9b: 1228-1233.
3. M. Ploeg, K.K.H. Aben, L.A. Kiemeny. 2009. The present and future burden of urinary bladder cancer in the world. *World J Urol* 27, 3: 289-93.
4. Macmillan cancer support. 2011. Symptoms of bladder cancer. Retrieved March 03, 2016 from <http://www.macmillan.org.uk/Cancerinformation/CancerTypes/Bladder/Symptomsdiagnosis/Symptoms.aspx>.
5. C. Zheng, Y. Lv, Q. Zhong, R. Wang, Q. Jiang. 2012. Narrow band imaging diagnosis of bladder cancer: systematic review and meta-analysis. *BJU International* 110: E680-E687.
6. A.C. James, J.L. Gore. 2013. The costs of non-muscle invasive bladder cancer. *Urol Clin North Am* 40, 2: 261-9.
7. K.D. Sievert et al. 2009. Economic aspects of bladder cancer: what are the benefits and costs? *World J Urol* 27, 3: 295-300.
8. C. Yeung, T. Dinh, J. Lee. 2014. The health economics of bladder cancer: An updated review of the published literature. *PharmacoEconomics* 32, 11: 1093-104.
9. NICE guidelines. 2015. Bladder cancer: diagnosis and management [NG2]. Retrieved March 03, 2016 from <https://www.nice.org.uk/guidance/ng2/chapter/1-Recommendations#diagnosing-and-staging-bladder-cancer-2>.
10. RCS. 2014. Good Surgical Practice. Retrieved March 3, 2016 from <https://www.rcseng.ac.uk/surgeons/surgical-standards/professionalism-surgery/gsp/documents/good-surgical-practice-pdf>.
11. Storz. Image 1: SPIES™. 2013. 96011020 TP2 SPIES 1.2 10/2013/MFL.E.
12. Mathias Lux, Michael Riegler. 2013. Annotation of endoscopic videos on mobile devices: a bottom-up approach. In *Proceedings of the 4th ACM Multimedia Systems Conference (MMSys '13)*, 141-145.
13. UNISOFT Medical Systems 2016. Retrieved March 03, 2016 from http://www.unisoftmedical.co.uk/urological_medicine.asp.
14. M. A. van Veelen, C.J. Snijders, E. van Leeuwen, R.H. Goossen, G. Kazemier. 2003. Improvement of foot pedals used during surgery based on new ergonomic guidelines. *Surg Endosc* 17, 7: 1086-91.
15. Matthias Wille, Sascha Wischniewski, Philipp M Scholl, Kristof Van Laerhoven. 2014. Comparing Google Glass with tablet-PC as guidance system for assembling tasks. In *Proceedings of the 11th International Conference on Wearable and Implantable Body Sensor Networks Workshops (BSN Workshops '14)*, 38-41.
16. R. Shadiev, N.S. W.-Y. Hwang, N.-S. Chen, Y.-M. Huang. 2014. Review of speech-to-text recognition technology for enhancing learning. *J Educ Techno Soc* 17, 4: 65-84.
17. G. Singh, A. Nelson, R. Robucci, C. Patel, N. Banerjee. 2015. Inviz: Low-power personalized gesture recognition using wearable textile capacitive sensor arrays. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom '15)*, 198-206.
18. Édimo Sousa Silva, Maria Andréia Formico Rodrigues. 2014. A Gesture Control System for Aiding Surgical Procedures. In *Proceedings of the 2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA '14)*, 287-296.
19. B. Brown-Clerk, J.B. Rousek, B.R. Lowndes, S.M. Eikhout, B.J. Balogh, M.S. Hallbeck. 2011. Assessment of electrosurgical hand controls integrated into a laparoscopic grasper. *Minim Invasive Ther Allied Technol* 20, 6: 321-8.
20. J.B. Rousek, B. Brown-Clerk, B.R. Lowndes, B.J. Balogh, M.S. Hallbeck. 2012. Optimizing integration of electrosurgical hand controls within a laparoscopic surgical tool. *Minim Invasive Ther Allied Technol* 21, 3: 222-33.
21. G. Harvin. 2014. Review of musculoskeletal injuries and prevention in the endoscopy practitioner. *J Clin Gastroenterol* 48, 7: 590-4.
22. J. Kim, X. Huo, M. Ghovanloo, 2010. Wireless control of smartphones with tongue motion using tongue drive assistive technology. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, 5250-5253.
23. L. Jing, Y. Zhou, Z. Cheng, T. Huang. 2012. Magic Ring: A finger-worn device for multiple appliances control using static finger gestures. *Sensors* 12, 5: 5775-5790.
24. Guilherme Cesar Soares Ruppert, Leonardo Oliveira Reis, Paulo Henrique Junqueira Amorim, Thiago Franco de Moraes, Jorge Vicente Lopes da Silva. 2012. Touchless gesture user interface for interactive image visualization in urological surgery. *World J Urol* 30, 5: 687-91.

25. David J. Ward, Alan F. Blackwell, David J. C. MacKay. 2000. Dasher—a data entry interface using continuous gestures and language models. In *Proceedings of the 13th annual ACM symposium on User interface software and technology*, 129-137.
26. Miika Silfverberg, I. Scott MacKenzie, Panu Korhonen. 2000. Predicting text entry speed on mobile phones. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '00)*, 9-16.
27. O.J. Muensterer, M. Lacher, C. Zoeller, M. Bronstein, J. Kubler. 2014. Google Glass in pediatric surgery: an exploratory study. *Int J Surg* 12, 4: 281-289.
28. B. A. Ponce et al. 2014. Emerging technology in surgical education: combining real-time augmented reality and wearable computing devices. *Orthopedics* 37, 11: (751-757).
29. P. M. Russell et al. 2014. First "glass" education: telementored cardiac ultrasonography using Google Glass- a pilot study. *Acad Emerg Med* 21, 11: 1297-1299.
30. Hui Min Knight, Parag Ravindra Gajendragadkar, Awais Bokhari. 2015. Wearable technology: using Google Glass as a teaching tool. *BMJ Case Rep* 2015: 1757-790
31. B. Benninger. 2015. Google Glass, ultrasound and palpation: The anatomy teacher of the future? *Clin. Anat* 28, 2: (152-155).
32. C.R. Davis, L.K. Rosenfield. 2015. Looking at plastic surgery through Google Glass: part 1. Systematic review of Google Glass evidence and the first plastic surgical procedures. *Plast Reconstr Surg* 135, 3: 918-28.
33. M.P. Schijven, M. Graafland, W.A. Bemelman. 2015. Google glass in surgery sharpen your vision. *Surg Endosc* 29, S72: 0930-2794.
34. P.R. Chai, K.M. Babu, E.W. Boyer. 2015. The feasibility and acceptability of Google Glass for teletoxicology consults. *J Med Toxicol* 11, 3: 283-287.
35. W. Vorraber et al. 2014. Medical applications of near-eye display devices: An exploratory study. *Int J Surg* 12, 12: 1266-1272.
36. G. Krishnamurthy. 2015. Google glass in intervention radiology-potential applications and limitations. *J Vasc Interv Radiol* 26, 2: 1051-0443.
37. H.M Mentis, A. Rahim, P.R. Theodore. 2015. Referencing ct scans through a headmounted optical display during laparoscopic surgery. *Surg Endosc* 29, S411: 0930-2794.
38. G. Aldaz et al. 2015. Hands-free image capture, data tagging and transfer using Google Glass: a pilot study for improved wound care management. *PloS one* 10, 4: e0121179, 1932-6203.
39. U.-V. Albrecht et al. 2014. Google Glass for documentation of medical findings: evaluation in forensic medicine. *J Med Internet Res* 16, 2: 1438-8871.
40. K. Horak, S.M. DeLand, D.S. Blair. 2014. The Feasibility of Mobile Computing for On-Site Inspection. *SAND* 2014: 18291.
41. M. Paterson, M.R. Glass. 2015. The world through Glass: developing novel methods with wearable computing for urban videographic research. *J Geogr Higher Educ* 39, 2 (Feb 2015), 275-287.
42. L. Mauerhoefer, P. Kawelke, I. Poliakov, P. Olivier, E. Foster. 2014. An exploration of the feasibility of using Google Glass for dietary assessment. Newcastle University, Computing Science, Technical Report Series, No. CS-TR-1419.
43. P. M. Scholl, Matthias Wille, Kristof Van Laerhoven. 2015. Wearables in the wet lab: a laboratory system for capturing and guiding experiments. In *Proc of Int Conf on Ubiquitous Computing*, 589-599.
44. P. M. Scholl, Tobias Schultes, Kristof Van Laerhoven. 2015. RFID-based compound identification in wet laboratories with google glass. In *Proc of Int Workshop on Sensor-based Activity Recognition and Interaction*, Article 13, 5 pages.
45. T. Wu, C. Dameff, J. Tully. 2014. Integrating Google Glass into simulation-based training: experiences and future directions. *J Biomed Graph Comput* 4, 2.
46. T. Yang, Y.M. Choi. 2015. Study on the design characteristics of head mounted displays (HMD) for use in guided repair and maintenance. In *Proc of Virtual, Augmented and Mixed Reality Conference*, 535-543.
47. J. M. Teixeira, R. Ferreira, M. Santos, V. Teichrieb. 2014. Teleoperation using Google Glass and AR, drone for structural inspection. In *Proceedings of Symposium on Virtual, Augmented and Mixed Reality*, 28-36.
48. R. McNaney et al. 2014. Exploring the acceptability of google glass as an everyday assistive device for people with parkinson's. In *Proceedings of CHI*, 2551-2554.
49. A. Khan et al. 2015. Beyond activity recognition: skill assessment from accelerometer data. *Proc. of Int Conf on Ubiquitous Computing*, 1155-1166.